

Checklista: Beständiga identifierare

Beständiga identifierare (persistent identifiers/PID:ar) är unika kodsträngar¹ kopplade till olika poster. De är en förutsättning för att man ska kunna skapa länkar i ett nätverk. Beständiga identifierare kan appliceras på både objekt och mer abstrakta konceptuella begrepp.

Varje kulturarvsinstitution bör ha en övergripande policy för användning av identifierare. Det är viktigt att identifierarna är utformade på ett sådant sätt att de är globalt unika och beständiga.² Ifall identifierarna endast används internt bör de vara utformade på ett sådant sätt att de kan integreras i formella standarder och format vid extern användning. Denna checklista beskriver olika identifierare närmare och vad din kulturarvsinstitution kan tänka på vad gäller beständiga identifierare.

Olika slags identifierare

Det finns olika typer av identifierare. Ett av de mer kända formaten är URI (Uniform Resource Identifier).³ En URI är en identifierare av en resurs på Internet eller i ett datornätverk och kan klassificeras som en plathänvisning (URL), ett namn för resursen (URN) eller både och.

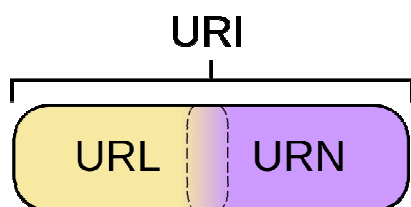


Bild 1. En URI kan vara en URL (Uniform Resource Locator)⁴, en URN (Uniform Resource Name) eller både och.

(Källa: "URI Euler Diagram no lone URIs" by David Torres original author derivative work: Qwerty0 (talk) - URI_Venn_Diagram.svg. Licensed under CC BY-SA 3.0 via Wikimedia Commons.⁵)

En beständig identifierare kan utgöras av en kombination av institutionens interna identifierare samt information om vilken typ av resurs som identifieras och hur resursen kan

¹ Siffror eller läsbara kodsträngar.

² Se mer information om vad som avses med detta i "Persistent Identifiers (PIDs): Recommendations for Institutions", ATHENA, <http://www.athenaeurope.org/getFile.php?id=779>, avsnitt 4, s. 29.

³ http://sv.wikipedia.org/wiki/Uniform_Resource_Identifier

⁴ http://en.wikipedia.org/wiki/Uniform_resource_name

⁵ http://commons.wikimedia.org/wiki/File:URI_Euler_Diagram_no_lone_URIs.svg#/media/File:URI_Euler_Diagram_no_lone_URIs.svg och http://en.wikipedia.org/wiki/Uniform_resource_name#/media/File:URI_Euler_Diagram_no_lone_URIs.svg

nås. Det finns även ett antal tjänster som gör det möjligt att länka mellan de identifierare som används internt i institutionernas samlingssystem och de identifierare som skapas när informationen t.ex. tillgängliggörs online. Med en uppslagstjänst (resolution service) kan man skapa beständiga identifierare av exempelvis interna identifierare vid institutionen, som identifierar en digital fil, men också referera till flera filer och digitala objekt eller lagra metadata om en resurs.

De flesta bibliotek använder idag URN:er (Uniform Resource Name) för publikationer, som genom en uppslagstjänst kan kopplas ihop med URL:er. En URL (Uniform Resource Locator) ger information om hur man når resursen och var den finns. På så sätt är det möjligt att göra skillnad mellan ett unikt namn och en plats, i de fall en namngiven resurs eller dess kopior finns på flera platser. Ibland kan det även vara relevant med specifika typer av identifierare, som t.ex. LSID:er som är identifierare för Life-Science.⁶

System för beständiga identifierare

Det finns både interna och externa system som kan hantera beständiga identifierare. Om man använder ett internt system ska man se till att utforma systemet så att det följer rekommenderade standarder för exempelvis utformning av URI:er. Om URI:erna skapas internt är det viktigt att de utformas så stabilt och beständigt som möjligt, vilket det står mer specifikt om i EU-kommissionens rekommendationer⁷ och i W3C:s rekommendationer⁸ om t.ex. ”cool URI”. Vid val av ett externt system är det viktigt att vara uppmärksam på att systemet är tekniskt tillförlitligt, auktoritativt, flexibelt i sättet att presentera metadata och interoperabelt.⁹ I relevanta fall kan även en uppslagstjänst som dirigerar webbläsaren till en aktuell kopia av objektet användas.

PID-system används i allt större utsträckning för digitala resurser eftersom URI:er generellt ses som otillräckliga på lång sikt pga. att domännamn och webbsidors struktur ändras över tid. URI:er är dock beständiga så länge en institution behåller registreringen av domännamnet. PID:ar skapade med ett PID-system är i sin tur beständiga så länge den databas som håller PID-registreringarna finns kvar. Med andra ord slutar inte PID:ar skapade med ett PID-system att vara beständiga för att institutionen t.ex. ändrar namnet på hemsidan. Samtidigt kräver ett PID-system tekniskt underhåll, vilket gör att PID:ar skapade i ett sådant system i praktiken kan vara mindre beständiga än domännamn eftersom skapandet av dem är beroende av kompetens och specialbyggd mjukvara.

⁶ http://www.tdwg.org/fileadmin/subgroups/guid/LSIDs_for_Biologists.pdf

⁷ https://joinup.ec.europa.eu/catalogue/asset_release/study-persistent-uris-identification-best-practices-and-recommendations-topi

⁸ <http://www.w3.org/TR/ld-glossary/#persistent-identifier-scheme>

⁹ Se mer information om vad som avses med detta i ”Persistent Identifiers (PIDs): Recommendations for Institutions”, ATHENA <http://www.athenaeurope.org/getFile.php?id=779>, avsnitt 5.

Med hjälp av en så kallad innehållsförhandling (content negotiation) kan olika typer av svar ges av servern beroende på vilken agent det är som efterfrågar URI:n, t.ex. kan en webbläsare få ett HTML-dokument eller en PDF-fil, en RDF-läsare får en RDF-fil och om det dyker upp nya format på vilka informationen kan representeras kan dessa läggas till senare.¹⁰

Beständiga identifierare vid din institution

- Finns det en datapolicy vid din institution?
- Finns det en policy för identifierare vid din institution?
- Är institutionens identifierare unika internt?
- Är institutionens identifierare unika globalt?
- Är institutionens identifierare beständiga internt?
- Är institutionens identifierare beständiga globalt?
- Om identifierarna endast används internt, är de utformade på ett sådant sätt att de kan integreras för extern användning på ett unikt och beständigt sätt, i enlighet med rådande rekommendationer?
- Är identifierarna som används redan implementerade i institutionens befintliga system (t.ex. i databaser som används vid registrerings- och/eller katalogiseringsprocesser)?
- Återanvänds de unika identifierare som redan finns utan att en ny identifierare skapas?
- Fungerar återsökbarheten (genom att det vid t.ex. uppdatering av länkar skapas en länk som pekar mot originalinformationen)?
- Finns det behov av en extern uppslagstjänst för identifierare (som t.ex. Kungliga bibliotekets uppslagstjänst¹¹)?
- Om de beständiga identifierarna uttrycks med URI:er, följer dessa följande standard rekommenderad av Europeiska kommissionen:
”http://{domän}/{typ}/{koncept}/{referens}”¹²

System för beständiga identifierare vid din institution

- Används systemet endast internt eller också externt?
- Är systemet integrerat med andra system vid institutionen, t.ex. samlingssystem?
- Finns det en plan för förvaltning och långsiktigt bevarande av systemet?

¹⁰ Råd gällande beständiga länkar utarbetade av Kommunförbundet Stockholms län, Peter Krantz AB, PwC och Sveriges kommuner och landsting med stöd av Vinnova,
http://skl.se/download/18.772895e214641fb368dcf000/1403190241384/Rad_gallande_bestandiga_lankar_SKL.pdf

¹¹ <http://www.kb.se/isbn-centralen/urnnbn/kallfil/>

¹² Enligt “Study on persistent URIs, with identification of best practices and recommendations on the topic for the MSs and the EC”, Europeiska kommissionen, 2012:
<https://joinup.ec.europa.eu/sites/default/files/D7.1.3%20-%20Study%20on%20persistent%20URIs.pdf>

- Finns det en introduktion till systemet för alla som använder identifierare vid registrering av material?
- Är systemet tekniskt tillförlitligt, t.ex. vad gäller:
 - säkerheten?
 - säkerhetskopiering med redundant teknik?
 - automatisk uppdatering av registret över identifierarna?
- Om det är ett externt system som används för att tilldela identifierare/resolvers till resurser, är systemet i sådana fall förvaltad av en auktoritativ och trovärdig organisation?
- Är systemet flexibelt och kan det exponera data utan att informationen förloras, dvs. sker exponeringen kvalitativt och med hög granularitet)?
- Är systemet interoperabelt (tillämpas t.ex. vedertagna internationella standarder och/eller öppna standarder)?

Exempel på standarder och system för beständiga identifierare

Uniform Resource Name (URN):

<http://tools.ietf.org/html/rfc3406>

URN:NBN (National Bibliography Numbers):

<http://www.kb.se/isbn-centralen/urnnbn/>

Persistent URL (PURL):

<https://purl.oclc.org/docs/index.html>

Digital object identifier (DOI):

<http://www.doi.org/>

Archival Resource Key (ARK):

<http://www.ifla.org/best-practice-for-national-bibliographic-agencies-in-a-digital-age/node/8793>

Open URL:

<http://www.oclc.org/research/activities/openurl.html>

Handle-systemet:

http://en.wikipedia.org/wiki/Handle_System

Persistent identifier component, som är en del av ECK (Europeana Connection Kit) och som i sin tur har utvecklats av Europeana Inside-projektet:

http://www.europeana-inside.eu/documents/communication_material/eck-guide.html

Best Current Practice RFC 7320 - URI Design and Ownership:

<https://tools.ietf.org/html/rfc7320>

Några fördjupningstips

”Persistent Identifiers (PIDs): Recommendations for Institutions”, ATHENA:

<http://www.athenaeurope.org/getFile.php?id=779>

”State of the art report on persistent identifier standards and management tools”, Linked Heritage, 2013:

<http://www.linkedheritage.eu/index.php?en/215/persistent-identifiers>

Learning object: ”Persistent identifiers – what if?”, Linked Heritage:

<http://linkedheritage.cab.unipd.it/training/LO-01/en/00-about.html>

Learning object: ”Persistent identifiers – commercial and heritage views”, Linked Heritage:

<http://linkedheritage.cab.unipd.it/training/LO-05/en/00.html>

”Persistent Identifiers for Cultural Heritage”, Digital Preservation Europe, Briefing Paper:

http://www.digitalpreservationeurope.eu/publications/briefs/persistent_identifiers.pdf

”Implementing Persistent Identifiers: Overview of concepts, Guidelines and Recommendations”, Consortium of European Research Libraries
European Commission on Preservation and Access, 2006:

<http://www.ica.org/5694/paag-resources/implementing-persistent-identifiers-overview-of-concepts-guidelines-and-recommendations.html>

”Study on persistent URIs, with identification of best practices and recommendations on the topic for the Member States and the European Commission”, Europeiska kommissionen, 2012:

https://joinup.ec.europa.eu/catalogue/asset_release/study-persistent-uris-identification-best-practices-and-recommendations-topi

En utökad webbaserad version:

<http://philarcher.org/diary/2013/uripersistence/>

”Persistent and unique identifiers”, CLARIN, 2008:

<http://www.clarin.eu/sites/default/files/wg2-2-pid-doc-v4.pdf>

The CIDOC Statement on LOD identifiers:

http://network.icom.museum/fileadmin/user_upload/minisites/cidoc/PDF/StatementOnLinkedDataIdentifiersForMuseumObjects.pdf

”MuseumID - Identifying museums and museum objects on the Internet”:

<http://museumid.net/documentation>

Några best practices från W3C-communityt (t.ex. ”URI construction”):

<http://www.w3.org/TR/2015/WD-dwbp-20150224/#dataIdentification>

<http://www.w3.org/TR/ld-bp/#HTTP-URIS>

Checklistan är utarbetad av Digisam i samråd med experter från Sverige och utlandet. Vi vill särskilt tacka Leif Johansson på SUNET, Gordon McKenna på Collections Trust och Antoine Isaac på Europeana.

Version 1.0. Licens CC-0. En senare version kan finnas på www.digisam.se

Digisam är ett samordningssekretariat för digitalisering, digitalt bevarande och digitalt tillgängliggörande av kulturarvet. Samordningssekretariatet är en av regeringen beslutad verksamhet vid Riksarkivet.